CURITY CLASSIFICATION OF THIS PAGE

AD-A252 235

REPORT DOCUMENTATION PAGE

DTIC ELECTE JUL 0 2 1992 S B D

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Essex Corporation; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Naval Air Systems Command
6c. ADDRESS (City, State, and ZIP Code) 1040 Woodcock Road Orlando, FL 32803	7b. ADDRESS (City, State, and ZIP Code) Washington, DC 20361
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Naval Air Systems Command; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00019-91-C-0149
8c. ADDRESS (City, State, and ZIP Code) Department of the Navy Washington, DC 20361	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
N00019	LQ220		

11. TITLE (Include Security Classification)
Correlating Visual Scene Elements with Simulator Sickness Incidence: Hardware and Software Development (U)

12. PERSONAL AUTHOR(S)
R. S. Kennedy, K. S. Berbaum, & M. G. Smith

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 4/91 TO 10/91	27 Oct 91	19

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
simulator sickness, flight simulators, visual displays, motion perception, human engineering, image generation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Simulator sickness occurs in a large number of Army, Navy, and Marine Corps simulators, and is most prevalent in moving-base, rotary-wing devices which employ cathode ray tube (CRT) video displays as opposed to fixed-wing, dome-display trainers with no motion base. Based on data from a factor analysis of over 1000 Navy and Marine Corps pilot simulation exposures, a new scoring procedure was applied to two helicopter simulators with similar rates of simulator sickness incidence. Based on the factor analytic scoring key, the two simulators showed slightly different sickness profiles. Preliminary work was begun to record the visual scene by video frame-by-frame decomposition and automated scoring algorithms were developed. The findings are discussed from the standpoints of (1) recommendations for future design and use of simulators, and (2) the metric advantages and other merits of the "field experiment" methodology to address human factors problems with simulator sickness.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL CDR Michael G. Lilienthal	22b. TELEPHONE (Include Area Code) (703) 692-1560; 22c. OFFICE SYMBOL PMA 205-3J1

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

92 7 01 023

92-17412

INTRODUCTION

Advancing technology in many areas regularly generates unwanted by-products, negative consequences that can become serious problems if they are not anticipated and resolved early in the developmental process. The world of simulator technology is no exception. Simulator sickness was first reported 30 years ago in two studies by Havron and Butler (1957) and Miller and Goodson (1960). Since that time, the number of studies and reports of simulator sickness has increased at an exponential rate; there was as much published on simulator sickness since 1988 as in all previous years. A simulator sickness R&D program, sponsored by the Naval Air Systems Command, which began in a formal way in 1982, initially emphasized problem definition. By 1984, once simulator sickness was determined to be a problem of some magnitude, the emphasis within the Navy's R&D effort was to establish guidelines to provide interim remedies for the Navy's simulators which were on-line. Since then, the emphasis has shifted to the identification of the nauseogenic properties of the stimulus, particularly the inertial forces and, to some extent, the visual characteristics of the stimulus.

The impact of simulator sickness to the Navy and Marine Corps is of four main types:

- o <u>Safety and Health</u>. There are verifiable increases in locomotor ataxia, interference with higher-order motor control, physiological discomfort, and visual aftereffects or flashbacks. These same skills are required for driving, flying or even roof repair and accidents may result from simulator exposure.

- o <u>Readiness and Operational Effectiveness</u>. Grounding policies are evolved when simulator sickness is reported because of its implications for safety and health. It has been estimated that the modal aviator's operational availability could be reduced by as much as 5% to 10% following published restrictions on flying in force at some facilities. Reduction of simulator sickness may result in very high payoff in improved operational readiness of Navy pilots.

- o <u>Acquisition Economics</u>. There are instances where: a) simulator systems have been specified and built with capabilities that cannot be used due to simulator sickness and b) where symptoms began to be reported AFTER a simulator was upgraded with wider field of view, greater luminance, a new visual etc. Program managers do not need these surprises.

- o <u>Training</u>. Negative opinions of simulator training may result when individuals are made sick during training. There may be negative transfer of learned skills from the simulator into actual control of the military system. It is possible that acquisition of habits inappropriate to control of operational systems may result from behaviors learned in simulators to avoid sickness.

Flight simulators provide an essential role for training in both military and commercial environments. They are utilized to train aircrew for almost every conceivable operational platform and their benefits in terms of cost,

safety, and flexibility of training assure that they will continue to be an integral part of military training (Orlansky & String, 1980). It has been estimated (Shelsby, 1989) that the simulation market, at about $2.49 billion in 1987, will be close to $6.23 billion in 1992. Accompanying the growth of simulation in training has been increasing sophistication in simulation technology, beginning with the Link "blue box" in the 1930s, and progressing to complex systems encompassing six-degree-of-freedom motion bases and detailed, near-photographic quality visual displays, visually-coupled virtual reality concepts, and mission rehearsal systems which may be used at sea.

Crucial to the design of simulators is specification of the equipment parameters that will promote training effectiveness and realism, but also avoid simulator sickness. However, the technological advances which have provided the opportunity for increased fidelity have, in turn, placed greater demands on other tolerances on simulator subsystems (e.g., responses of visual and motion base systems and their interaction). Visual display systems combine diverse methodologies for generating and enhancing visual information, and sometimes through misalignment, failure, or other factors, eyestrain and other symptoms related to motion sickness may be experienced. Yet test pilots may be unaware of the source of these difficulties and are therefore sometimes unable to provide enough information for the visual display engineer to identify and correct the problem. Needless to say, standards and specifications to address these problems are also lacking. In consequence, effective training may be compromised, and flight simulator subsystems may be purchased that cannot be used, and so the Navy does not get good value for their acquisition dollars.

Despite the best intentions of program managers, project engineers and simulator developers, newly fielded or upgraded systems are often accompanied by unexpected pilot reports of simulator-induced discomfort. To develop specifications that will avoid this problem, methods of recording and analyzing the visual display must be developed so as to relate the visual stimulus with desired and undesired vision-based responses. Standards can only be created when such data are available. This solution sounds simple, but it is not. To be maximally effective, not only must the entire visual and inertial stimuli presented to the pilot be recorded, but also what the pilot "sees" and "controls" must be extracted from a mere record of the physical energy per se. Distinguishing what is effective input to the human, as well as the part of the input that is self-generated via operator control may be crucial in determining causes of the problems reported by test pilots and experienced aircrew. Fortunately, recording the visual stimulus such stimuli is becoming increasingly simple with television cameras. The plan for this Phase I feasibility demonstration was to follow twc parallel measurement approaches.

The first technical step was to measure the problem(s) experienced by the pilots as accurately as possible. The causes cannot be determined until there is a suitable assessment of the "criterion." The criterion against which the engineering characteristics will ultimately be compared needs to be reliable and valid and sufficiently diverse so that differential stimulus effects can be discriminated. Considerable experience and success have been shown with the use self-report scales of the Cooper-Harper type, but which were specifically developed to address the perceptual discontinuities experienced

by pilots (Fowlkes, Kennedy, & Allgood, 1990). More than 5000 simulator exposures have been analyzed and reported from Navy (Fowlkes et al. 1990; Kennedy, Lilienthal, Berbaum, Baltzley, & McCauley, 1988) and from Marine Corps (Kennedy & Smith, 1990), and Army (Gower, Lilienthal, Kennedy, Fowlkes, & Baltzley, 1987) flight simulators.

The second step was to record the stimulus (the physical visually displayed scene) and analyze it in the same way as the various "filters" in the human that interpret and respond to the stimulus. To the extent that our image processing mimics the humans filters through which the physical presentation must pass, the results of our analyses should match the incidence of symptoms. We know that the human visual-vestibular systems are very responsive to motion of several different types (viz., loom, vection, displacement, velocity, etc). The human nervous system responds to "higher order" stimulus structure that interacts with the motor control outputs of the pilot. The prototype reported on here is a first step toward a more ambitious and complete analyses of the stimulus-response relationships that underlay sickness in simulators. Eventually, in the next Phase of this project, the pilot's eye movements (which provide the best evidence available for where the pilot is looking -- although not necessarily what the pilot "saw") must eventually be recorded and related through analysis to the replayed visual scene. In addition, and perhaps equally important, the pilot's control (i.e., stick, throttle, rudder, collective, etc.) must also be recorded and related to aspects of the visual scene because these effector controls influence human perceptions, although their influence is not well known outside the neuroscience laboratories (cf. e.g., Burnham & Aertker, 1970; Lackner & DiZio, 1984; Lackner & Graybiel, 1981; Leibowitz, Post, & Sheehy, 1986; Mather & Lackner, 1981; Post & Leibowitz, 1985; Reason & Benson, 1978; von Holst, 1954). Because the human efferent motor control system influences how the human perceives his/her inputs, failure to include these elements will reduce the chance for eliminating this problem (Kornhuber, 1974; Mayne, 1974).

While the neurophysiology and photomechanical properties of the human sensory system are reasonably well known (cf., Boff, Kaufman & Thomas, 1986; Vol I, II & III), the scientific literature does not provide sufficient guidance to write specifications to avoid simulator sickness when visual and inertial inputs are presented to humans in combination in ways that do not occur in physical reality. When physical fidelity is not achieved by a single stimulus modality, training effectiveness can be compromised. If multiple inputs are presented that are in poor agreement with themselves and with the past history of an experienced pilot, simulator sickness can result. Similar problems are being experienced by the National Aeronautics and Space Administration in connection with the Space Adaptation Syndrome (i.e., space sickness) and a major scientific program is ongoing to examine and remedy these problems at Ames Research Center, Moffett Field, CA, and Johnson Space Center, Houston, TX.

A major obstacle is the lack of basic and applied research on how the human visual and vestibular motion perception systems code and filter. Much simulation research has focused on image generation and enhancement, but only relative to the static aspects of visual displays and sufficient basic data are not available on linkages and interactions of visual and inertial perceptions of motion so that simulation parameters can be adjusted when

sickness problems arise. The DOD has not funded 6.1 or 6.2 research that would lead to information from which such specifications could be written. This omission has been repeatedly brought home when Navy and Air Force sponsored 6.3 experimental research has been conducted using small samples of experienced pilots in high-cost state-of-the-art experimental simulators. Carefully conducted "measure everything" empirical studies at government facilities at NASA, Ames; Williams AFB; and Visual Technology Research Simulator (VTRS) at the Naval Training Systems Center have yields inconclusive findings. A "white paper" which explores many of the statistical and methodological difficulties has been sponsored by PMA205 (Kennedy, & Fowlkes, 1991). This study concludes that the best research opportunity for circumventing these problems is to use simulator sites (OFT, WST, etc.) where actual aircrew regularly receive training as field experimental laboratories. However, in order to conduct experiments, the "stimulus" must be known. Scientific data recording capability at these sites are not easily available. Extensive modification and intrusion on operational readiness has often been required for engineers to carry their standard laboratory recording equipment, place it in a simulator and make measurements. Additionally, a 2-hour hop can result in mountains of recorded data and all of these physical measures must then be related to one or more measures of the pilot's experience. We think that it is methodologically and statistically important not to attempt to measure everything. Guidance provided by perceptual theory, coupled with some technological opportunities which have come available within the past two years make the present approach feasible. First, factor analysis of self-reports by pilots and aircrew of simulator exposure (Kennedy, Lilienthal, Berbaum, Baltzley & McCauley, 1988; Kennedy, Smith, & Jones, 1991; Kennedy, Fowlkes, & Lilienthal, 1991) which are described in greater detail below, *imply that* if the pilots' report (which are reliably different in different simulators) could be linked to the engineering features of the simulator which produced them. Second, high-speed image processing capability at relatively low cost have become available. The merger of these two technologies was the goal of this effort.

<u>Background of Phase I and II Technical Objectives</u>

We proposed to conduct two separate development efforts in parallel:

- o An automated recording method was to be assembled in which pilot output in the form of self-report would be employed to evaluate the effects of visual scenes, particularly in those cases where test pilots and aircrew might provide their assessment of flight simulator systems but are unaware of (or cannot describe) the source of the problem. Essex scientists have been developing motion sickness tests of all kinds for Army, Navy and NASA for more than 20 years and the self report form was originally developed >30 years ago by Dr. Robert S. Kennedy and Dr. Ashton Graybiel at the Navy's Aerospace Medical Research Laboratory.

- o A transportable recorder and image processing system for recording characteristics of the visual scene and what the pilot sees and controls. Essex (Orlando Office) was particularly well-suited to conduct this development through its 10-year history of investigations into the visual aspects of flight simulators at the Naval Training Systems Center's Visual Technology Research Simulator (Kennedy, Fowlkes, Berbaum, & Lane, 1989), as well as unique experience in the field of signal and image processing.

Development of the on-site recording of the visual scene was an empirical effort which would take place at the simulator sites. The development of the scoring techniques for the motion sickness symptom questionnaires was analytic and made use of data obtained in the Navy's 10 simulator study or data recently collected by Essex.

DEVELOPMENT OF THE AUTOMATED VISUAL SCENE RECORDING SYSTEM

Video Recording Methods

Our expressed purpose was to record actual man-in-the-loop simulations using video recorders at two different flight simulators and develop ways of computer-analyzing that video data so that the visual scenery could be decomposed into meaningful psychophysical cues. Technical aspects of importing this video data and creating a programming environment within which to perform analyses were important technical subgoals. The development of algorithms to automatically score the strength of stimuli believed to be important for the experience of self-motion in simulators (interactions with the ground plane, eyeheight, tilt, roll, loom) in terms of position, velocity, and acceleration indicants could then begin. Finally, we could prepare to relate these measures of the visual stimulus for self-motion to the different symptom profiles based on instances of self-reports of sickness in samples of pilots.

The first recording was of the 2F121 simulator at Marine Corps Air Station, New River, NC, a simulator for the CH-53E helicopter. The recordings were made during the afternoon while a qualified helicopter pilot flew the simulation through a series of prescribed settings and motions. For this preliminary analysis, a standard NATOPS-specified syllabus hop was performed, and no special constraints were imposed on the pilot. Much of the time was spent following the runway, turning to landing, and flying at low and intermediate altitude. The second recording was of the 2B42 at NAS Whiting Field, FL., a simulator for the TH-57C helicopter. Again the recordings were made during the afternoon while a qualified helicopter pilot flew the simulation through similar maneuvers as were flown as in the 2F121.

Both video recordings were made with a hand-held Panasonic portable mini-cam, which was stationed immediately behind the pilot. The lens was set to manual focus at 9 feet. Particular attention was paid to recording the visual scene within the central 20 degrees around the forward heads-up line of sight from the operator's head position because it is within that limited cone wherein most eye movements and visual perceptions take place (Sanders, 1970). In Phase II, the visual scene and the pilot output systems would be integrated so that they would be automatically recorded, scored and correlated. In addition, head and eye movements will be recorded and stick inputs will be monitored and analytically decomposed so that what a pilot "sees" and "feels," as well as what the pilot "controls," is integrated during playback into a single prototype system. Further, the inertial motion could also be linked to the visual scene.

Image Analysis Methods

Data were analyzed using a video frame grabber (Redlake Corporation, Spectrum NTSC+) and various image processing analyses have been completed. For our initial study, we attempted to characterize the most important, substantial, and general attributes of the visual display likely to be related to simulator sickness. A measure highly related to altitude is important because of the role of eyeheight in determining the impact of textures on self motion (Warren, 1982). The horizon was selected as a primary target for image analysis because of its usefulness and reliability as an indicant of numerous types of important motion for pilots. The horizon can be used as an index of pitch attitude and altitude if its change is taken into account. That is, while a single position of the horizon in the field of view on any frame might result for various combinations of altitude and attitude, rapid rates of change across frames can only result from pitch variation. (It may be necessary in the Phase II device to analyze gradients of motion in the ground plan to differentiate altitude and pitch change in an exact way.) The instantaneous horizon value was determined every second. This computation took the form of an area measure in order to disentangle it from the angle of the horizon. Change in horizon position (velocity) and change in velocity (acceleration) were calculated over 1-minute epochs; therefore, the sample size for each 1-minute data point is 60. These 60-element 1-second epochs were then plotted over a uniform hop length; in this case, 5 minutes. Thus, the horizon vertical position was the first variable, rate of position change was the second variable, and rate of change in the rate of position change was the third variable.

Variables four, five and six were obtained by performing a similar analysis in which the left and right edges of the video were algebraically summed and compared with the horizon position on the left and right edges of the frame. In this way the angle of roll (left wing down/right wing up or right wing down/left wing up) can be determined. This computation was also performed 60 times a minute, and this comprised the roll position. Additionally, the position was integrated over each 60-second period so that roll velocity and acceleration were obtained. These 1-minute intervals were plotted for an entire 5-hop session length. Variables seven, eight, and nine represented an experimental attempt to determine movement in depth of various objects in the scene in order to provide a measure comparable to the first six for the loom (that is, monocular movement in depth). We believed that some of these measures would have a high correspondence to visual stimulus for perceived self motion (which is also linked to eyestrain via oculomotor driving) and therefore to the magnitude of the sickness and type of symptoms experienced.

This initial project is crucial for the plan of later phases which will relate differences in symptom profile between simulators to differences in indices of the visual stimulus. For Phase I, we only attempt to establish sensitivity of the analysis to visually depicted kinematics. This outcome would suggest that our measures are sensitive to the types of kinematics performed in these devices and suggest thereby that we are able to capture these fundamental aspects of the visual stimulus. Also, if any differences could be observed between our indices of the visual displays of the two simulators with the same maneuvers performed, then we might hope to relate

them to differences in symptom profile experienced by pilots. This much evidence obtained so early in development, though not a completely adequate demonstration of its feasibility, lends some credence to this approach.

DEVELOPMENT OF SYMPTOM PROFILE SCORING TECHNOLOGY

Historically, motion sickness scientists employ motion sickness symptomatology questionnaires (Kennedy, Tolhurst, & Graybiel, 1965) to handle the problem of different symptoms being experienced by individuals. The MSQ reflects the polysymptomatic nature of simulator sickness in that multiple symptoms are taken into account in the diagnostic scoring.

The theory behind scaling motion sickness severity is that vomiting, the cardinal sign of motion sickness, is ordinarily preceded by a combination of symptoms (Lentz & Guedry, 1978; McNally & Stuart, 1942; Money, 1970). Therefore, in order to score motion sickness beyond merely a vomit/no-vomit dichotomy, Wendt (1968) initially employed a three-point continuum scale in a series of studies on motion sickness. This scale was used to assess motion sickness symptomatology, whereby vomiting was rated higher than "nausea without vomiting" which, in turn, was rated higher than discomfort. Navy scientists developed a Motion Sickness Questionnaire (MSQ) consisting of a checklist of symptoms ordinarily associated with motion sickness for use in sea and air sickness studies (Kennedy, Tolhurst, & Graybiel, 1965). These symptoms included: cerebral (e.g., headache), gastrointestinal (e.g., nausea, burping, emesis), psychological (e.g., anxiety, depression, apathy), and other less characteristic indicants of motion sickness such as "fullness of the head." A response was required for each symptom using a rating of "none," "slight," "moderate," or "severe" (or in some cases "yes" or "no"). From this checklist, a diagnostic scoring procedure was applied resulting in a single, a five-point symptomatology scale, serving as a global score reflecting overall discomfort. The five point scale was expanded in studies of seasickness conducted by the U. S. Coast Guard, with the cooperation of the U.S. Navy, (Wiker & Pepper, 1978; Wiker et al., 1979a, b; Wiker, Pepper, & McCauley, 1981).

These scoring techniques are useful in that they permit quantitative analyses and comparisons of motion sickness in different conditions, exposures, and environments. However, a deficiency for the study of simulator sickness is that the single global score does not reveal information about the potentially separable dimensions of simulator sickness and it lacked statistical normalization properties. It was argued that such information could be informative about the nature of simulator sickness and may also serve a diagnostic function; not just about the individual but to signal differences in the equipment factors (e.g., visual distortion; motion characteristics) which may differentially cause the sickness.

Simulator Sickness Questionnaire (SSQ)

In order to obtain information about the separable dimensions of simulator sickness, >1000 Motion Sickness Questionnaires (MSQ) have been factor analyzed (Lane & Kennedy, 1988). The results of that study produced three specific factors and one general factor. The three factors form the basis for three SSQ subscales. These subscales or dimensions appear to operate through

different "target" systems in the human to produce undesirable symptoms. Scores on the Nausea (N) subscale are based on the report of symptoms which relate to gastrointestinal distress such as nausea, stomach awareness, salivation, and burping. Scores on the Visuomotor (V) subscale reflect the report of oculomotor-related symptoms such as eyestrain, difficulty focusing, blurred vision, and headache. Scores on the Disorientation (D) subscale are related to vestibular disarrangement such as dizziness and vertigo. It was also found that one could abbreviate the list of symptoms to 16 items and this would result in little loss in accuracy. Subsequently, a Simulator Sickness Questionnaire (SSQ) was developed based on these 16 symptoms only. In addition to the three subscales, an overall Total Severity (TS) score, similar in meaning to the old MSQ score, is obtained. Each SSQ subscale was scaled to have a zero point and a standard deviation of 15.

When compared to the distribution of symptoms for other forms of motion sickness simulator sickness shows pronounced visual problems (e.g., eyestrain) whereas neurovegetative (nausea) phenomena predominate in seasickness. In our experience, the profiles of symptomatology from these two environments would be expected to show the widest disparity.

The scores which are used to evaluate the performance of the simulator can be arithmetic means but ordinarily the incidence of sickness is decidedly non-normal. This means, for example, that, should the average be 10, there may be as many as 50 percent of the pilots with essentially no sickness and 20 percent with scores above 50. Thus, while it is always desirable for simulators to have overall lower scores, even a simulator with a low score may still place some pilots at risk after leaving the simulator (Kennedy, Lilienthal, Berbaum, Baltzley, & McCauley, 1989). Since conducting the work reported above, we have turned to the use of an additional score to index the "safety" issues of a simulator and we have given some thought to the score attained by the 75th percentile person. In our view, anyone with a score higher than 20 should be warned of his/her condition and not permitted to leave the simulator building unless extreme care is used. Anyone with a score over 15 should contact a flight surgeon or corpsman or be carefully debriefed by an experienced instructor pilot.

EFFECTIVENESS OF SYMPTOM PROFILE AND VISUAL SCENE RECORDING

Two separate objectives were achieved. First, an automated recording method was assembled in which human output was employed to evaluate the effect of visual scenes. Second, a transportable recorder and image processing system for recording engineering characteristics of the visual scene of what the pilot sees and controls was undertaken.

Automated Profile Scoring of Simulator Sickness

Figure 1 shows simulator sickness data from several simulators within the Navy and Marine Corps inventory and may be used to construct the "Fleet Average". It may be seen that differing amounts of sickness are obtained in the different simulators and the helicopter devices with a moving base have the highest incidence. As described above, this approach is based on the development of the Simulator Sickness Questionnaire (SSQ) (Lane & Kennedy, 1988). It should be recalled that these central tendency scores may be

employed to rate the simulator for incidence severity but also to provide "subscale" scores which are considered to be more diagnostic of the locus of simulator sickness in a particular simulator for which overall severity is shown to be a problem. By "sharpening" the measuring instrument, the ability to do "differential diagnoses" on simulators may allow better identification and evaluation of engineering solutions to problems and lead ultimately to more precise specifications for simulator design and guidelines for simulator use.

In the simulators considered for the present study, using Total Scores, five helicopter simulators are compared for overall incidence in figure 1. Shown here are two models of the 2F121 (one in MCAS New River, NC and one in MCAS Tustin, CA); one model of the 2F120 (simulating the CH-53D - a helicopter with pilot population, mission and simulator characteristics seemingly similar to the CH53-E) and two models of the 2B52 (simulating the primary helicopter flight trainer - TH57). In addition to the data reduction of the pilot reports of sickness, for the present report, the 2F121 and the 2B52 were visited by Essex scientists and hops were video recorded in each and returned to the Essex Orlando laboratory for development of the image processing techniques.

Whereas the total score measures are shown in Figure 1 for these five simulators, Figure 2 displays the three factor solution employed in our more recent work for these same five simulators. It may be seen that the symptomatology in these three models of simulators (2F121, 2F120, 2B52) may be decomposed into three separable dimensions. Note that there is total comparability (indeed, overlap) of the symptom profile between the "sister" models and note further that there are decided differences between models. Furthermore, in Figure 2, the profile of sickness may be seen to be different from the impression one obtains when looking at Figure 1. That is, the relative mixture of symptomatology in the five simulators may be seen to have the following characteristics: (1) nausea as a symptom cluster is highest in the 2B42 simulator and eyestrain is the least. Conversely, eyestrain is highest in the 2F121 and lowest in the 2B42. It may be seen that disorientation is also highest in the 2F121 and somewhat lower in the 2B42. The lowest disorientation score is in the 2F120. The 2B42 and the 2F121 may be seen to be markedly different when compared to each other, although the two models of the 2B42 appear to exhibit a profile of symptomatology which is exactly the same between models 2 and 4 of that simulator. Likewise, the 2F121, New River, and 2F121, Tustin, simulators which reside more than 3,000 miles apart, show almost exactly the same profile of symptomatology.

In conclusion we believe that this method of scoring shows considerable promise, not only for separating simulators according to total score incidence (figure 2) but also for determining if one of the symptom clusters is showing a higher than average incidence. It could be argued that when a higher than average incidence in a particular cluster is seen, this can provide suggestions as to which of the several equipment features is likely to be the source of the problem.

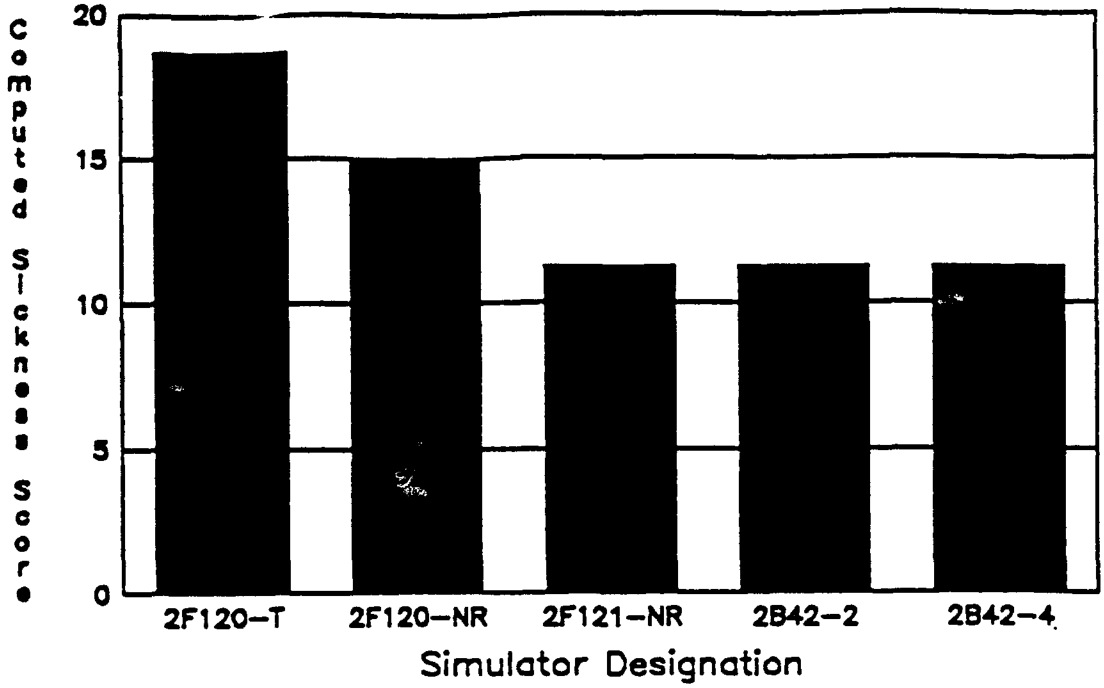

Figure 1. Total sickness score across five simulators.

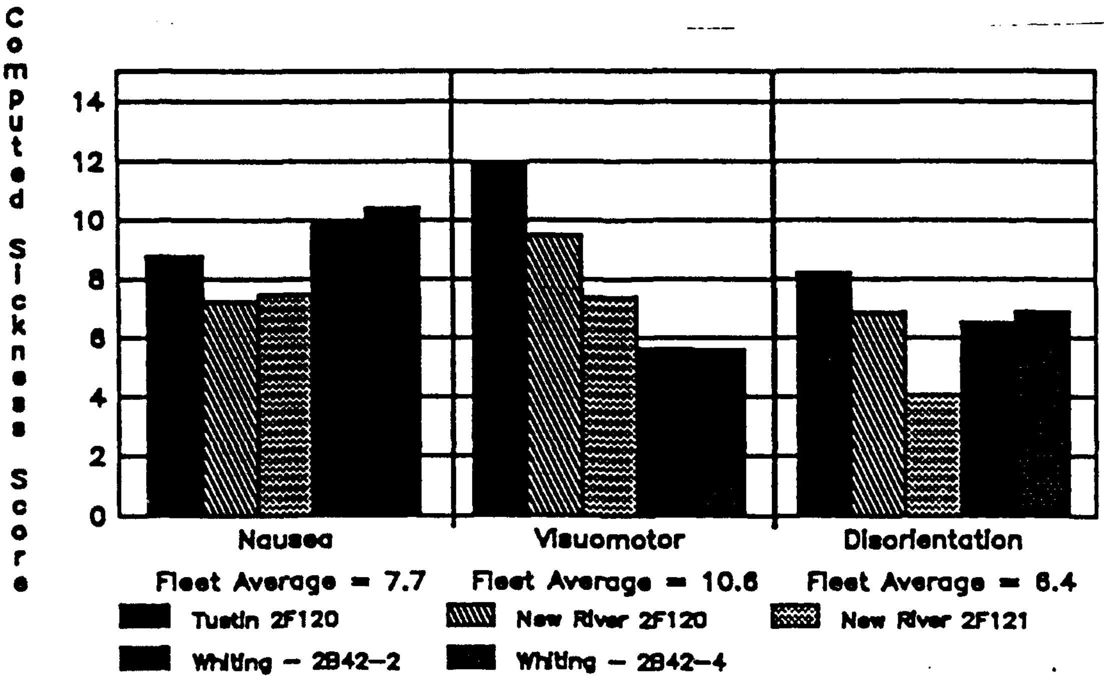

Figure 2. Spectral composition of simulator sickness across five simulators.

Transportable Recorder and Frame-by-Frame Analysis

At the outset of this Phase I R&D effort, the scale of the units of analysis that would be most useful was not known. We worried whether our initial globally directed approach might have been too gross. We now believe that these concerns were unfounded. It appears that our video recordings are very effective in capturing visual display data in the field and that our global approach to analysis may also be quite effective. For example, because the magnitude of the vection (self-motion perception) experienced can be expected to be predictive of the magnitude of the sickness (Hettinger, Berbaum, Kennedy, Dunlap, & Nolan, 1990), we needed some method to assess vection from the visual display. We initially focused on the 20 degrees of visual angle at the center of the display as this encompasses most fixations. It seems that sampling this part of the visual display is very informative with respect to trajectory and will represent many components of the analysis of larger areas. We have also focused on computerized inspection of the visual display for tracking the position of the horizon. This was selected since close interaction with the ground plane is one of the most potent predictors of simulator sickness (Kennedy & Fowlkes, 1991). In this regard, we have succeeded in developing what appear to be very robust algorithms for horizon position locating and tracking. Moreover, the horizon is visible in 97% of the frames we have analyzed, and by application of expectancy based on history of horizon movement, we are successful in predicting the location of horizon reappearance. Analysis of horizon angle and height in the field provides several crucial kinds of information: angle provides tilt position and integration of position with respect to time provides roll velocity. An "activity" measure which integrated the number of changes in direction can be expected to provide additional information regarding strength of the stimulus from the standpoint of simulator sickness and can be used in conjunction with the other metrics related to roll. Similarly, changes in height would be descriptive of texture and altitude. Additional visual features under investigation include visual display area with relative motion (parallax) of visual elements, average rate of these motions, and density of elements participating in the motion. These elements, taken together, should provide an integrated aggregate of the vection stimulus (perception of apparent self-motion) which has been implicated as a primary factor (cf., Hettinger et al., 1990) in simulator sickness (particularly in fixed-base simulators).

It had been decided that we would record the visual as a pilot flew a series of maneuvers in two flight simulators in order to determine whether differences could be obtained through automatic recording and subsequent analyses by a videotape of the scene from the pilot's eye point in both devices. The plan would then be to relate differences in the two systems to profile measures of the two systems as evidenced by the Automated Symptom Profile Measurement Approach.

Figure 3 may be seen to contain a series of photocopies of the visual scene as obtained through the frame grabber analysis. The horizon may be seen to provide a very useful angle across the video that can provide an index of the amount of roll instantaneously in time and when integrated the amount of roll over time. Tilt and roll were expected to be useful activity measures for predicting simulator sickness symptomatology.

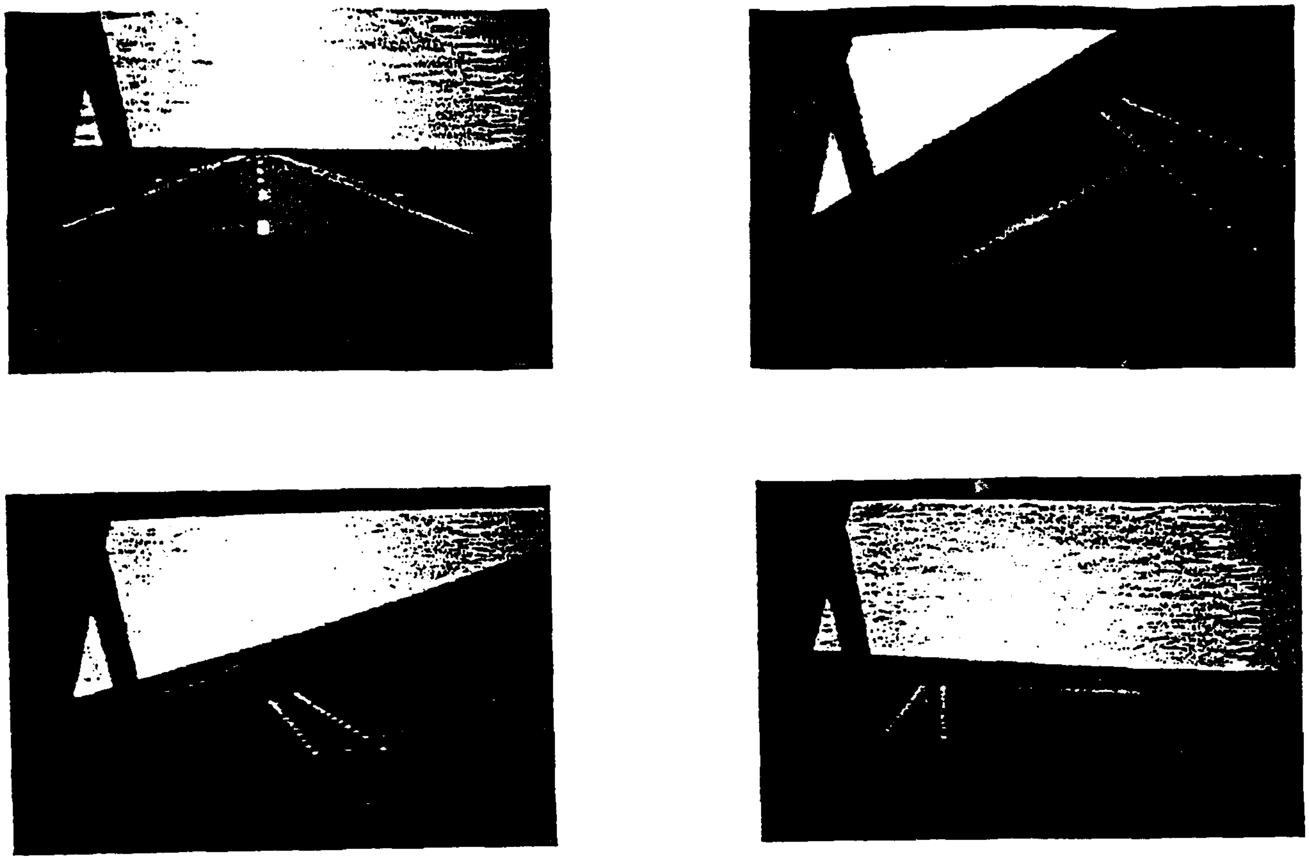

Figure 3. Images captured from videotape of TH57-C.

It is difficult to appreciate from still figures, how effectively these algorithms track key display elements. To better illustrate our current level of performance, we have included a videotape and a computer program (running on a 386) that show horizon tracking of simulator displays recorded during actual naneuvers. Values of several of our display parameters are simultaneously displayed. That this analysis can be performed using microcomputer technology bodes well for the routine analysis of visual display systems.

DISCUSSION OF OUTCOMES AND PROSPECTS FOR PHASE II

Because the ultimate criterion for acceptance of flight simulator visual systems is filtered through the human visual system, we considered that the first task to be performed was to develop a suitable evaluation system based on that filter. Ordinarily, such approaches are for purposes of aeromedical management, and the data acquisition process is to determine whether the pilot is adversely effected (is he/she sick?). However, the same data acquisition system can be employed in order to make estimates about whether changes have occurred in the visual system or other engineering aspects of the equipment. In this case, the question is whether the flight simulator (or its visual components) is sick. Pilots who exhibit extreme reactions are at risk for simulator-induced posteffects and need to be identified so they can be warned or restricted with regard to post-training activities. The second question pertains to quality assurance (QA) testing of simulators and is most pertinent to the present proposal. Based on pilots' reactions (i.e., symptomatology), the level of simulator sickness and other visually-induced problems can be tracked over time for a given simulator using a "quality control" model to

detect shifts in calibration or other gradually emerging problems. This requires routine on-site data collection in a near "automated" mode.

Simulator sickness has been an attendant consequence of simulation, even from its earliest days, and there is evidence that increasing sophistication in simulation systems may aggravate the problem (Kennedy, Allgood, & Lilienthal, 1989). Symptoms, which include nausea, stomach awareness, disorientation, eyestrain, and headache, have been reported by instructors and students training in all forms of vehicular simulation (automobile, aircraft, tank), but have been most extensively documented in military flight simulators. Simulator sickness may be caused by visual imagery which is improperly designed ("out-of-spec") or from changes which occur over time due to ordinary usage, lack of maintenance, misalignment, etc. Below, we briefly describe examples of engineering measurements which may be candidates for the problems which pilots observe when flying the flight simulator systems, but which they may not be able to verbalize as to cause.

We have been surprised by the level of our success in the initial feasibility study of our approach to characterizing the visual display. We have been able to demonstrate that capture of the relevant visual data is straightforward using even the simplest of modern video technology, and that we can capture a very large sample of measurements within missions or hops which can then be integrated into an overall measure of visual impact upon the vestibulo-ocular systems. This constitutes a fundamental "proof of concept" that routine visual display capture and analysis of simulator is practical. This measurement occurs with the human-in-the-loop and measures aspects of the displays that can only be estimated with other approaches. It promises a level of validation that cannot be achieved in any other way. Of course, funding our Phase II proposal will be necessary to bring these procedures to a mature form. We believe that this approach will succeed in capturing another large element of the simulator sickness problem which will thereby lead to its eventual elimination.

We would argue that these results demonstrate sensitivity of the technique to different maneuvers within a single simulator and to simulators with different symptom profiles flying similar maneuvers. Therefore, we will have a firm foundation on which to design a device that captures even more complex aspects of display (such as distortion, misalignment, control-display lag, and cue asyncrony) and we can then attempt a detailed characterization of simulator visual displays. The central reason for studying sensitivity to magnitude of stimulation first is that all other factors are likely to interact with stimulus strength. (Misalignment would have very little impact on vection-induced disorientation and eyestrain if the strength of the stimulus for vection is low.) The simulator equipment factors most likely to contribute to reduced fidelity, pilot problems, and simulator sickness are: (1) the presence of static and dynamic visual distortions, (2) improper calibration of out-the-window visual displays, (3) temporal lags between control inputs and subsequent visual display and/or motion base updates, (4) the presence of asynchronous relations between visual and vestibular cuing.

Visual Distortions. In the proper calibration of light infinity optics there must be proper calibration by application of computerized ray-trace analysis (Ebenholtz, 1988). Also, Rosinski (1982) made the important point

that graphic displays, such as those used in flight simulator visual systems, provide accurate representations of three-dimensional space only when viewed from the geometric center of projection. If the head is moved outside the center of projection, geometric distortions occur in the projected imagery which provide inappropriate visual information for self-motion. The presence of these distortions may account for the findings reported by Crosby and Kennedy (1982) that aircrew who viewed a simulator's visual display from an off-axis position experienced greater symptomatology. Under static conditions, up to about 20 diopters, the greater the off-axis distortion, the greater the disruption and the longer the adaptation (Dewar, 1970).

The deleterious effect of optical distortions may be magnified with highly detailed imagery, since irregularities are likely to be more noticeable as the amount of visual information is increased. The use of wide-field-of-view visual systems may also magnify the effects of distortions. The peripheral visual system is particularly sensitive to motion stimulation, and the irregularities in motion patterns introduced by distortions may provide inappropriate self-motion information.

Alignment of Visual Display Channels. There are two primary reasons for checking the alignment of computer-generated imagery (CGI) channels. First, misalignment means that there is no design eye for the complete system from which all channels can be viewed simultaneously. Thus, the same distortions that result from having one's head outside the design eye could occur with misaligned CGI channels. Second, if CGI optical channels had different foci, then a scene depicted at infinity would require different accommodative distances. The resulting extensive accommodative search could result in fatigue and headaches. There is evidence that this occurs in regular visual display terminal usage (deGroot & Kamphuis, 1983; Shahnavaz & Hedman, 1984; Smith, Tanaka, & Halperin, 1984). Simulators which are suspected of being out of alignment should be evaluated. Disparate base sites of adjacent CGI displays may also be disruptive.

Visual and Inertial Lags. Computational limitations generally produce temporal lags between operator control inputs and subsequent changes in position as indicated by the visual display and motion base. Seevers and Makinney (1979) expressed doubt concerning whether the motion system cuing scheme of the SAAC contributed to simulator effectiveness. Their evaluation of motion-base responses disclosed excessive lag times and cross-coupling between movements, indicating the existence of errors in the movement of the platform in relation to pilot control inputs. Inaccuracies in motion cuing such as those described by Seevers and Makinney (1979) have been thought to contribute to simulator sickness and to disruptions in perceptual-motor performance. Thus far, three experiments have addressed the issue of simulator sickness as related to lags and asynchronies. In the first study (Uliano, Kennedy, & Lambert, 1986), three asynchronous visual throughput delays were investigated in a fixed-base simulator. There were no differences in sickness rates among the three stimulus conditions. In the second study (Frank, Casali, & Wierwille, 1987), transport delays were shown to have an effect on performance (i.e., manual control) behaviors, but simulator sickness incidence did not appear to be related to the size of the delay. The third study (Hettinger & McCauley, 1989), from which only preliminary data are available, contrasted a fixed-base condition with nominal motion base

parameters, increased lead (resulting in attenuated motion), and a reduced motion bandwidth condition. Preliminary data, based on mean discomfort ratings, show little difference between the fixed-base and nominal conditions, and slightly elevated ratings for the increased lead and reduced bandwidth conditions.

Most modern flight trainers employ CGI. Computational operations at a rate of 30 Hz require about 33 msec to generate an updated image, although it has been suggested that phase shifts of less than 30 degrees to 45 degrees at 1 Hz (83-125 msec) probably will not affect the control of simulated flight (Ricard & Puig, 1975). Nearly all the research dealing with computational lags in flight simulator visual displays is concerned with performance deficits as a function of delay. Whether certain delays are more or less conducive to simulator sickness has not been extensively studied (Uliano et al., 1986). It may be that performance deficits and physical discomfort follow different functional relationships relative to the magnitude of delay. It is known that visual-motion lags may produce pilot-induced oscillations which may have two consequences: (1) produce nauseogenic very-low-frequency simulator motion, and (2) produce dynamic visual distortions because of the load imposed on the computer system.

Puig (1970) pointed out that optimal lag time is probably not a constant but may be a function of the intensity of the stimulus, as indicated by much of the work of K. U. Smith (1963) on delayed perceptual feedback. Howard and Templeton (1966) have questioned Smith's results, although it is fairly well accepted that temporally and/or spatially displaced sensory feedback generally impedes learning and disrupts performance. The magnitude of the delay which degrades motor performance may not be the same value as the interval which one might find most distressing. Both of these forms of delay are present in flight simulators, but generally only the delay which intrudes on performance is studied.

In Phase II, we will propose a larger scale comparison of the human performance methodology against on-line engineering monitoring of simulator systems. An on-line diagnostic system for simulator performance monitoring is needed to detect changes in system output (i.e., running out of tolerance) that may compromise training. We will evaluate two approaches to the on-line monitoring of fielded systems. The first approach utilizes human output (i.e., symptomatology or other responses) to monitor systems performance. Simulator sickness occurs because humans are sensitive to sensory conflicts and other provocative aspects of simulation. Thus, for example, if a simulator runs out of tolerance on some factor that increases sensory conflict, the measurement of symptomatology reported by individuals who use the system should reflect this change. Thus, human performance (i.e., rating of symptomatology) may be used to monitor systems performance. The second approach entails obtaining on-line engineering measures to monitor systems performance. These may include motion base tolerances, lags, asynchronies, and parameters of the visual system. It is our belief that, because of cost and reliability factors, human performance monitoring may offer a supplement or alternative to on-line engineering measures of systems. The method was favorably evaluated in terms of reliability, diagnostic capability,

REFERENCES

Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds.). (1986a). Handbook of perception and human performance: Volume I. Sensory processes and perception. New York: Wiley.

Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds.). (1986b). Handbook of perception and human performance: Volume II. Cognition and performance. New York: Wiley.

Burnham, C. A., & Aertker, D. R., Jr. (1970). Rotary motion and efferent readiness. Perception & Psychophysics, 7(5), 311.

Crosby, T. N., & Kennedy, R. S. (1982). Postural disequilibrium and simulator sickness following flights in a P3-C operational flight trainer. Bal Harbor, FL: Paper presented at the 53rd Annual Aerospace Medical Association.

deGroot, J. P., & Kamphuis, A. (1983). Eyestrain in VDU users: Physical correlates and long-term effects. Human Factors, 25, 409.

degroot, J. P., & Kamphuis, A. (1983). Eyestrain in VDU users: Physical correlates and long-term effects. Human Factors, 24, 409.

Dewar, R. (1970). Adaptation to displaced vision: Amount of optical displacement and practice. Perception & Psychophysics, 8(5A), 313-316.

Ebenholtz, S. M. (1988). Sources of asthenopia in Navy flight simulators. Columbus, OH: Battelle Memorial Institute. (NTIS No. AD A212 699)

Fowlkes, J. E., Kennedy, R. S., & Allgood, G. O. (1990). Biomedical evaluation and systems-engineering for simulators (BESS). Paper presented at the International Training Equipment Conference and Exhibition (ITEC). Birmingham, England.

Frank, L. H., Casali, J. G., & Wierwille, W. W. (1987). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. Proceedings of the Human Factors Society (p. 492). Santa Monica, CA: Human Factors Society.

Gower, D. W., Jr., Lilienthal, M. G., Kennedy, R. S., Fowlkes, J. E., Baltzley, D. R. (1987). Simulator sickness in the AH-64 Apache combat mission simulator (USAARL 88-1). Fort Rucker: U.S. Army Aeromedical Research Laboratory.

Havron, H. D., & Butler, L. F. (1957). Evaluation of training effectiveness of the 2FH2 helicopter flight trainer research tool (NAVTRADEVCEN 1915-00-1). Port Washington, NY: Naval Training Device Center.

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., & Nolan, M. D. (1990). Vection and simulator sickness. Military Psychology, 2(3), 171-181.

Hettinger, L. J., & McCauley, M. E. (1989). Visual motion synchrony in flight simulation. Paper presented at PM TRADE Simulator Sickness Briefing, Orlando, FL, December.

Howard, I. P., & Templeton, W. B. (1966). Human spatial orientation. London: Wiley & Sons.

Kennedy, R. S., Allgood, G. O., & Lilienthal, M. G. (1989). Simulator sickness on the increase. Paper presented at the AIAA Flight Simulation Technologies Conference, Boston, MA.

Kennedy, R. S., & Fowlkes, J. E. (1991, in press). Simulator sickness is polygenic and polysymptomatic: Implications for research. International Journal of Aviation Psychology.

Kennedy, R. S., Fowlkes, J. E., Berbaum, K. S., & Lane, N. E. (1989). Review of human performance research at the Visual Technology Research Simulator (NTSC TR89-020). Orlando, FL: Naval Training Systems Center.

Kennedy, R. S., Fowlkes, J. E., & Lilienthal, M. G. (1991, in press). Postural and psychomotor performance changes in Navy and Marine Corps pilots following exposures to flight simulators.

Kennedy, R. S., Lilienthal, M. G., Berbaum, K. S., Baltzley, D. R., & McCauley, M. E. (1988). Symptomatology of simulator sickness in 10 U.S. Navy flight simulators (NTSC-TR-87-008). Orlando, FL: Naval Training Systems Center.

Kennedy, R. S., & Smith, M. G. (1990). Simulator sickness in two Marine Corps helicopter trainers: Influence of configuration, usage, and pilot history (Final Rep., Contract DAAL03-86-D-0001). Research Triangle Park, NC: Battelle.

Kennedy, R. S., Smith, M. G., & Jones, S. A. (1991, April). Variables affecting simulator sickness: Report of a semi-automatic scoring system. Paper presented at the 6th International Symposium on Aviation Psychology, Columbus, OH.

Kennedy, R. S., Tolhurst, G. C., & Graybiel, A. (1965). The effects of visual deprivation on adaptation to a rotating environment (NSAM 918). Pensacola, FL: Naval School of Aerospace Medicine.

Kornhuber, H. H. (Ed.). (1974). Vestibular system part 2: Psychophysics, applied aspects and general interpretations. New York: Springer-Verlag.

Lackner, J. R., & DiZio, P. (1984). Some efferent and samotosensory influences on body orientation and oculomotor control. In L. Spillman & B. W. Wooten (Eds.), Sensory experience, adaptation, and perception. Hillsdale, NJ: Erlbaum.

Lackner, J. R., & Graybiel, A. (1981). Illusions of postural, visual, and aircraft motion elicited by deep knee bends in the increased gravitoinertial force phase of parabolic flight. Experimental Brain Research, 44, 312-316.

Lane, N. E., & Kennedy, R. S. (1988). A New method for quantifying simulator sickness: Development and application of the simulator sickness questionnaire (SSQ) (Technical Report EOTR 88-7). Orlando, FL: Essex Corporation.

Leibowitz, H. W., Post, R. B., & Sheehy, J. B. (1986). Efference perceived movement and illusory displacement. Acta Psychologica, 63, 23-24.

Lentz, J. M., & Guedry, F. E. (1978). Motion sickness susceptibility: A retrospective comparison of laboratory tests. Aviation, Space, and Environmental Medicine, 49(11), 1281-1288.

Mather, J. A., & Lackner, J. R. (1981). Adaptation to visual displacement: Contribution of proprioceptive, visual, and attentional factors. Perception, 10, 367-374.

Mayne, R. (1974). A systems concept of the vestibular organs. In H. H. Kornhuber (Ed.), Vestibular system part 2: Psychophysics, applied aspects and general interpretations (pp. 494-580). New York: Springer-Verlag.

McNally, W. J., & Stuart, E. A. (1942). Physiology of the labyrinth reviewed in relation to seasickness and other forms of motion sickness. War Medicine, 2, 683-771.

Miller, J. W., & Goodson, J. E. (1960). Motion sickness in a helicopter simulator. Aerospace Medicine, 31, 204-212.

Money, K. E. (1970). Motion sickness. Psychological Reviews, 50(1), 1-39.

Orlansky, J., & String, J. (1980). Reaping the benefits of flight simulation. Defense Management Journal, 16, 6-13.

Post, R. B., & Leibowitz, H. W. (1985). A revised analysis of the role of efference in motion perception. Perception, 14, 631-643.

Puig, J. A. (1970). Motion in flight training: A human factors view (NAVTRADEVCEN IH-177). Orlando, FL: Naval Training Systems Center.

Reason, J. T., & Benson, A. J. (1978). Voluntary movement control and adaptation to cross-coupled stimulation. Aviation, Space, and Environmental Medicine, 49(11), 1275-1280.

Ricard, G. L., & Puig, J. A. (1975). Delay of visual feedback in aircraft simulators. Applied Psychology, 59, 2250.

Rosinski, R. R. (1982). Effect of projective distortions on perception of graphic displays (Tech. Rep. No. 82-1). Pittsburgh, PA: University of Pittsburgh, Office of Sponsored Research.

Sanders, A. F. (1970). Some aspects of the selective process in the functional visual field. Ergonomics, 13(1), 101-117.

Seevers, J. A., & Makinney, R. L. (1979). Simulator for air-to-air combat motion system investigation (AFHRL-TR-79-18). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.

Shahnavaz, H., & Hedman, L. (1984). Visual accommodation changes in VDU operators related to environmental lighting and screen quality. Ergonomics, 279, 1071.

Shelsby, T. (1989). Military turning to simulators to cut training costs. The Sun, August 13.

Smith, A. B., Tanaka, S., & Halperin, W. (1984). Correlates of ocular and somatic symptoms among video display terminal users. Human Factors, 26, 143.

Smith, K. U. (1963). Special review: Sensory feedback analysis in medical research. I. Delayed sensory feedback in behavior and neural function. American Journal of Physiology, 42, 49.

Uliano, K. C., Kennedy, R. S., & Lambert, E. Y. (1986). Asynchronous visual delays and the development of simulator sickness. Proceedings of the 30th Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. Journal of Animal Behavior, 2, 89-94.

Warren, R., & Owen, D.H. (1982). Functional optical invariants: A new methodology for aviation research. Aviation, Space, and Environmental Medicine, 53(10), 977-983.

Wendt, G. R. (1968). Experiences with research on motion sickness (NASA Special Publication No. SP-187). Pensacola, FL: Fourth Symposium on the Role of Vestibular Organs in Space Exploration.

Wiker, S. F., Kennedy, R. S., McCauley, M. E., & Pepper, R. L. (1979). Reliability, validity, and application of an improved scale for assessment of motion sickness severity (USCG Tech. Rep. No. CG-D-29-79). Washington, DC: U.S. Coast Guard Office of Research and Development.

Wiker, S. F., Kennedy, R. S., McCauley, R. E., & Pepper, R. L. (1979). Susceptibility of seasickness: Influence on hull design and steaming direction. Aviation, Space, and Environmental Medicine, 50, 1046-1051.

Wiker, S. F., Pepper, R. L., & McCauley, M. E. (1981). A vessel class comparison of physiological, affective state, and psychomotor performance at sea. In J. C. Guignard & M. M. Harbeson (Eds.), Proceedings of the International Workshop on Research Methods in Human Motion and Vibration Studies. New Orleans: Naval Biodynamics Laboratory.